

(12) UK Patent Application (19) GB (11) 2 082 820 A

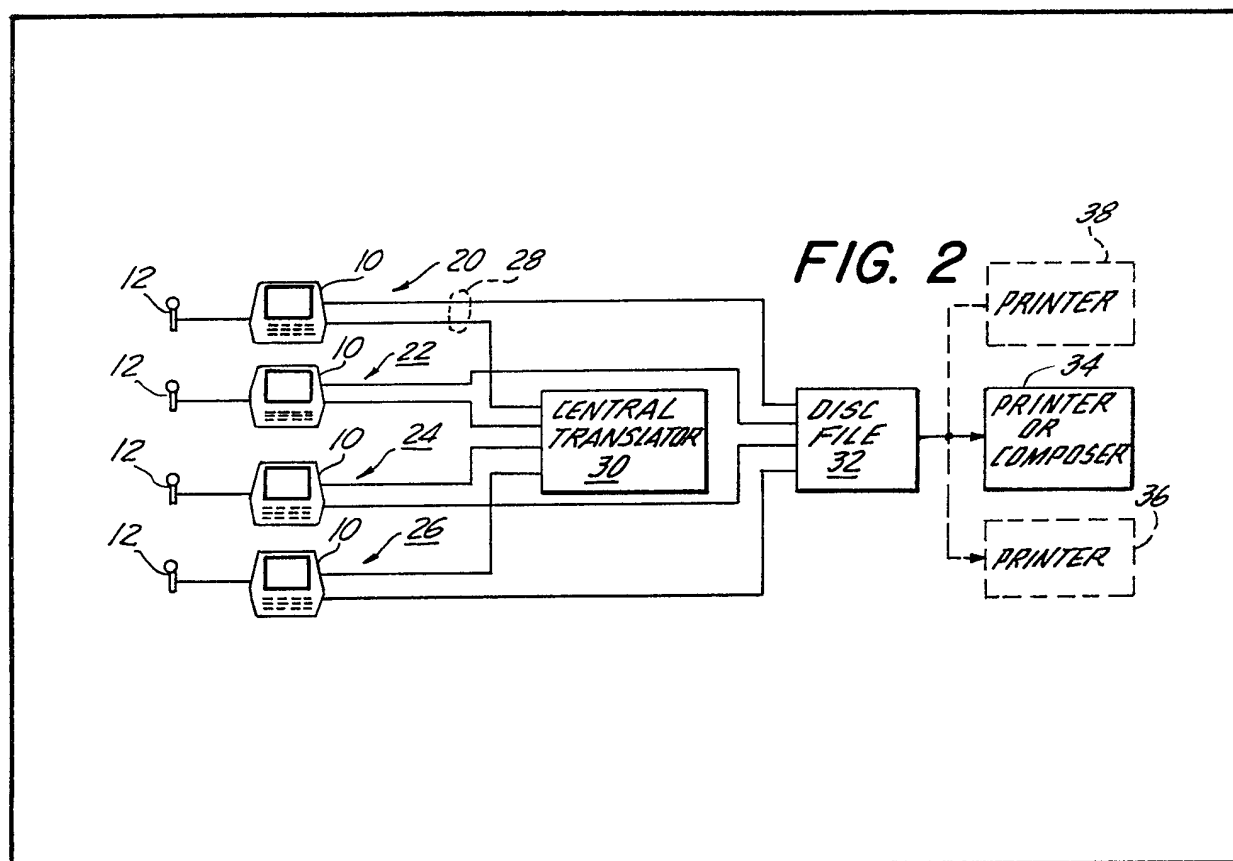
(21) Application No **8125511**
 (22) Date of filing **20 Aug 1981**
 (30) Priority data
 (31) **179808**
 (32) **20 Aug 1980**
 (31) **220108**
 (32) **24 Dec 1980**
 (33) **United States of America (US)**
 (43) Application published **10 Mar 1982**
 (51) **INT CL³**
G06F 3/16 G10L
1/00//A61B 5/00
 (52) Domestic classification
G4R 10B 10X 11D 11E
12A 12G 1F 3E 3G 3X 9B
9C
 (56) Documents cited
GB 2033633A
GB 1569450
GB 1418958
GB 1208262
GB 1039580
GB 978303
WO 79/00382A
 (58) Field of search
G4R

(71) Applicant
Gregor Nicholas Neff
85, Myrtle Avenue, Dobbs
Ferry, New York 10522,
United States of America
 (72) Inventor
Gregor Nicholas Neff
 (74) Agents
Baron & Warren,
16, Kensington Square,
London W8 5HL

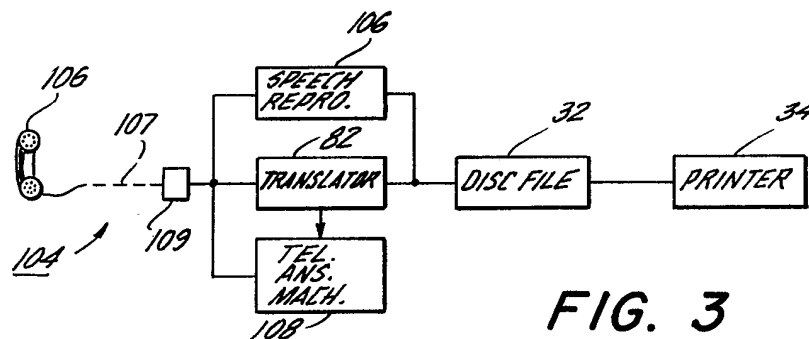
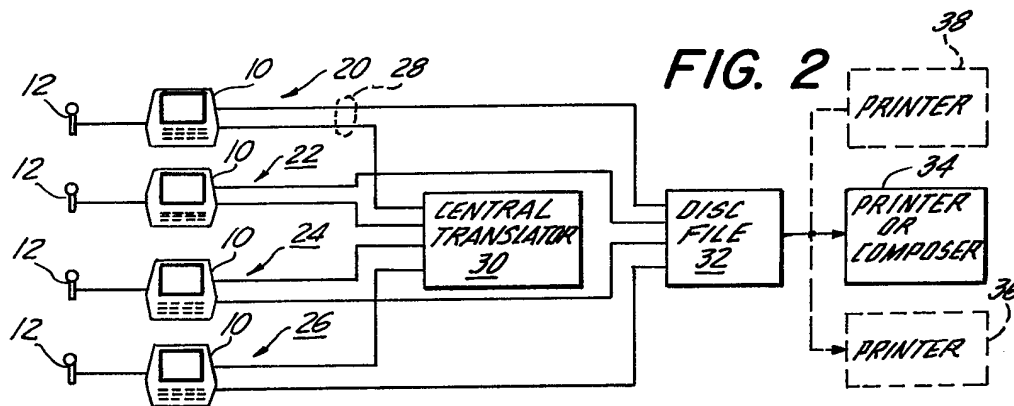
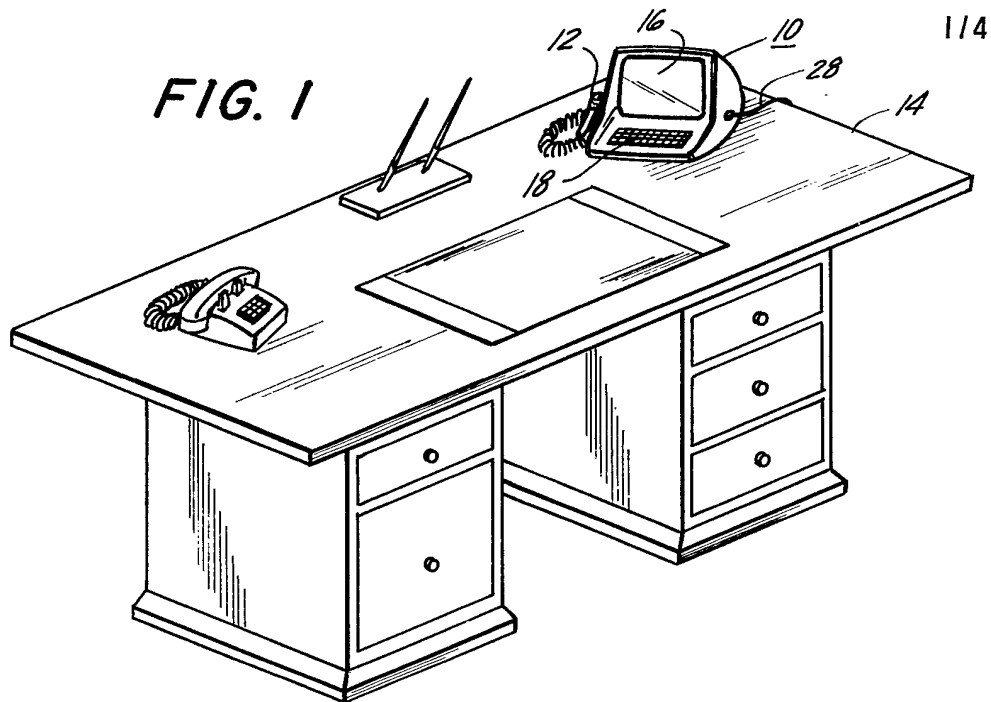
(54) **Devices, Systems and Methods for Converting Speech into Corresponding Written Form**

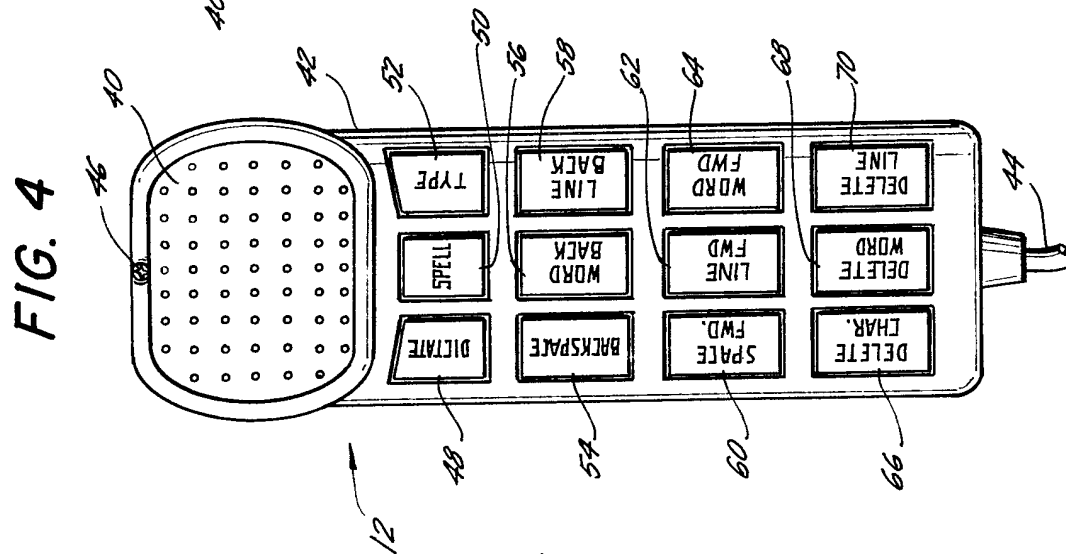
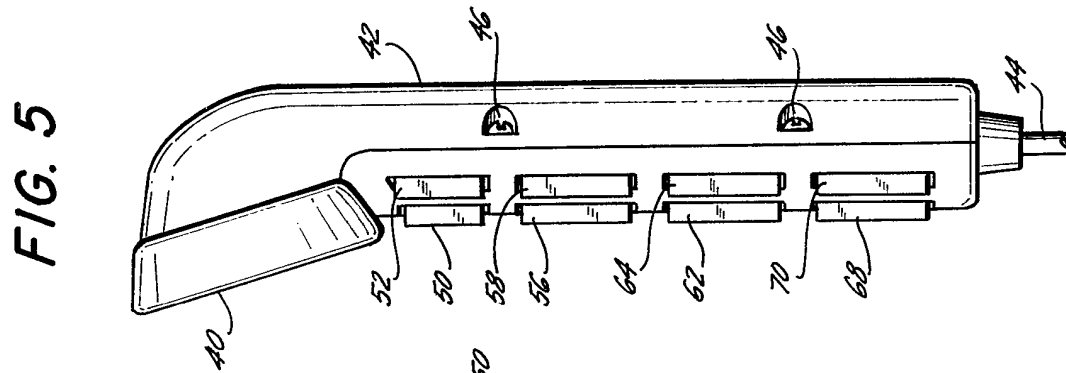
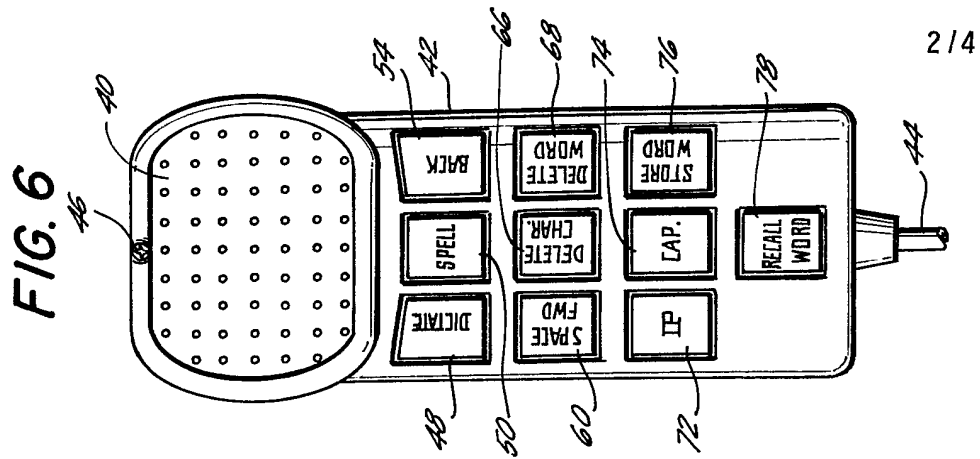
(57) In a system for converting speech into corresponding written words, the spoken word is converted directly into the written word by the use of an automatic speech recognition device or "translator" (30), which translates the spoken word into corresponding coded signals, and a word processing system including a dictation terminal (10). To compensate for the limited

vocabulary and other shortcomings of the translator (30), special control means are provided for allowing the dictator to verbally spell out each word or expression which the translator is incapable of translating; alternatively such words may be entered from a keyboard. In this mode of operation, the system automatically assembles the letters together to form words, and assembles the words so formed with other words to form sentences. The spoken words and letters appear on the video display screen of the word processor dictation terminal (10). The dictator views the text, makes any necessary corrections, and either stores the text codes in an electrical storage device (32) for later transmission to a printer, or directly transmits the text codes to a printer (34, 36, 38) which prints the text as a relatively high speed. The spoken words may be recorded on magnetic tape which is then read at a speed suitable for the translator (30).



GB 2 082 820 A





3/4

FIG. 7

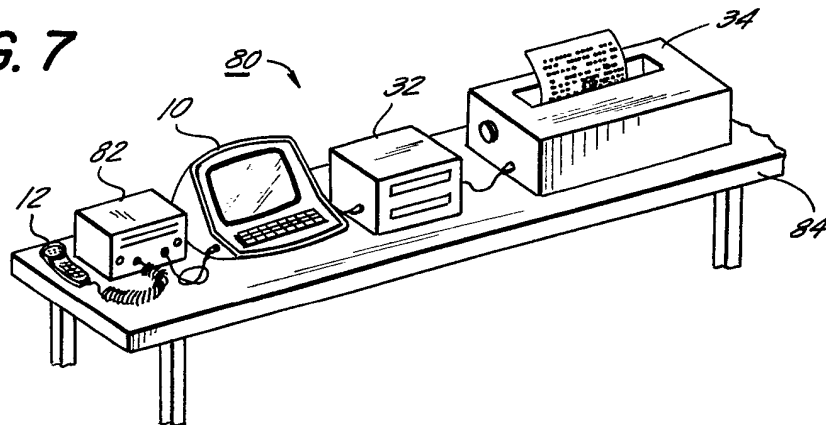


FIG. 8

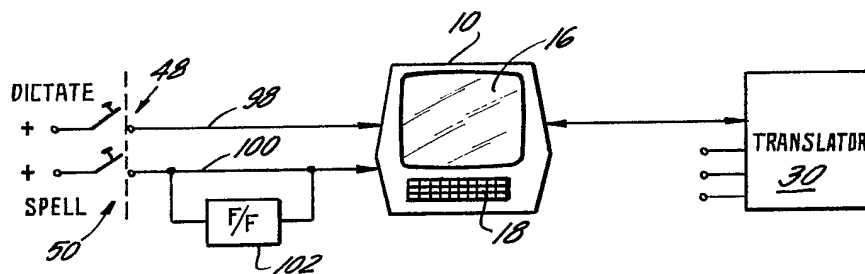
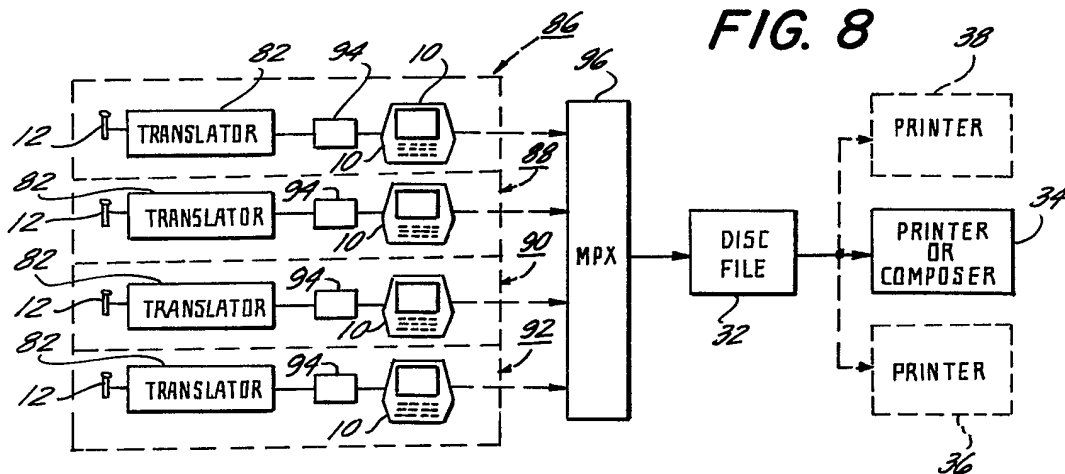


FIG. 9

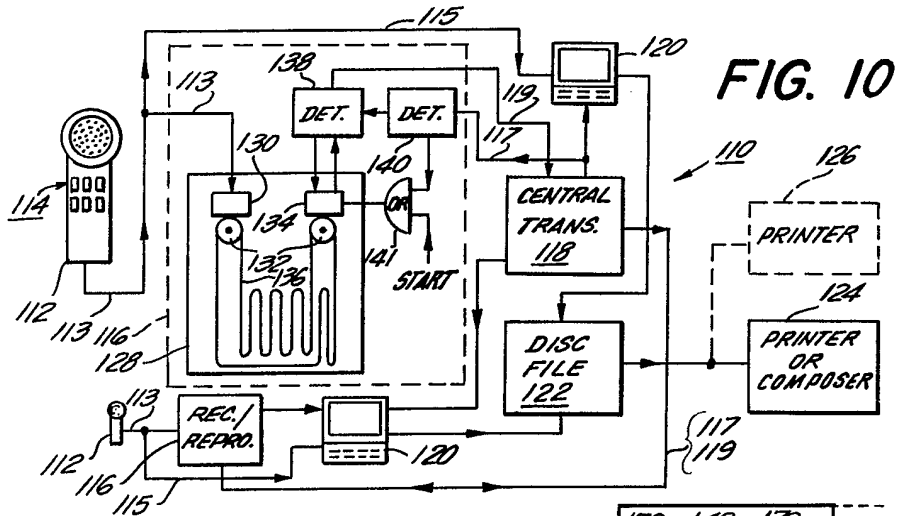


FIG. 10

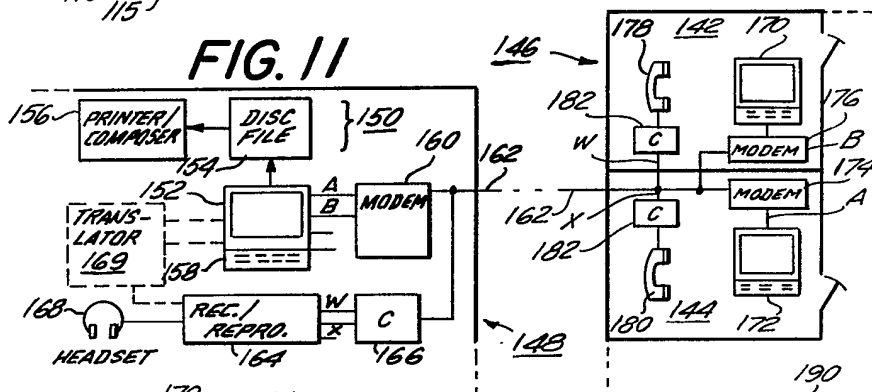


FIG. 11

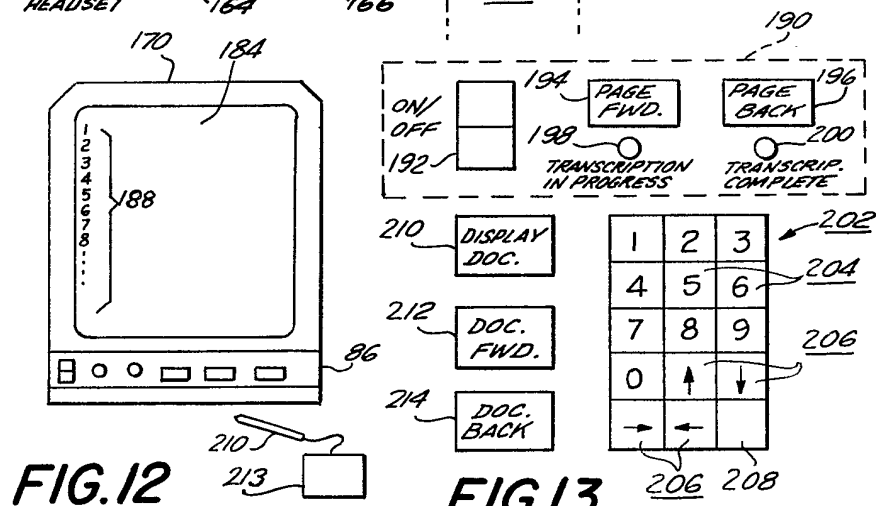


FIG. 12

FIG.13

SPECIFICATION

Devices, Systems and Methods for Converting Speech Into Corresponding Written Form

This invention relates to devices and methods for converting the spoken word into the written word; more specifically, this invention relates to devices and methods for dictating, transcribing and recording dictation in written form without human intervention between the dictation and the recording steps.

It long has been desired to provide a machine which will convert the spoken word directly into the written word. To this end, a substantial amount of work has been done on automatic speech recognition devices or "translators". Such devices convert spoken words or characters into coded electrical signals, which then can be displayed, printed or otherwise utilized. If such translators were perfect, it would be a relatively simple matter to utilize them in the automatic printing or typing of speech. However, such devices are quite far from perfect.

One drawback of the translators presently available commercially is that they have a relatively limited vocabulary. Most such translators have vocabularies of from thirty to three hundred words, and the most sophisticated machines claim to have vocabularies of from one thousand to two thousand words. This, of course, is unsatisfactory for use in most dictation since a dictating machine should be capable of handling virtually any word, character, symbol, or numeral in the language used by the dictator.

Another problem with translators presently available is that they are incapable of satisfactorily handling the problems caused by homonyms; that is, words which sound the same but have different spellings (e.g., "see" and "sea" are homonyms, as are "bear" and "bare"). The translator is not capable of discerning the proper spelling of the word from the text, and thus may not spell the word correctly when translating it.

A similar problem with translators presently available is that they usually require the uneconomical use of programming and memory capacity to handle the translation of the proper names of persons or places.

Another problem with the available translators is the cost. The cost usually is directly proportional to the size of the vocabulary of the machine, as well as its speed of operation. It is believed, therefore, that the cost of a translator with the size of vocabulary required for reasonably complete dictation capabilities would be prohibitive.

An additional problem with many translators is that the translator must be programmed to recognize words spoken by a particular individual. In doing the programming, usually the individual must speak each individual word into the machine several times in order for the machine to properly record a word recognition pattern against which the machine can compare words spoken later by the individual. Of course, the larger the vocabulary

of the machine, the longer the time it takes to properly program it.

One drawback with most conventional dictating systems is that the dictator himself does not promptly see a visible representation of the dictation and thus cannot review, edit and correct the text until later, after it has been transcribed.

Accordingly, it is an object of the present invention to provide an automatic or semi-automatic device and method for converting speech directly into written words; that is, to convert the spoken word into printed or typed form. More particularly, it is an object of the invention to provide such a device and method in which the problems and shortcomings mentioned above have been alleviated or eliminated.

The devices and methods embodying the invention enable the effective use of commercially available translators with relatively limited vocabularies. They enable homonyms and errors of the translator to be corrected relatively quickly and easily, either in advance, or shortly after dictation, without additional handling or personnel. Furthermore, they enable proper names to be handled relatively efficiently and accurately. Individual programming time for each individual using the equipment is minimized. The device and method are relatively simple and low in cost.

In accordance with one aspect of the present invention, there is provided a dictation writing device using a translator for translating the spoken word into electrically coded form, and a printer responsive to the coded signals for printing the words. At the dictator's option, the translator can be specially adapted to enable the operator to orally spell words which are incapable of being translated by the translator mechanism. The machine then correctly assembles the word or words formed by the spelling technique together with translated words in order to form sentences. The printer then prints the words so formed.

Preferably, the translator is used together with a word processing system having a video display terminal with a keyboard. The dictator sees the words he has dictated almost immediately as they appear on the screen of the terminal. If the machine has made an error in translation, or if the dictator finds that the word is untranslatable, or if he knows in advance that the word is untranslatable, he merely switches the machine into the "spell" mode of operation and spells out each word until the problem has been solved.

Preferably, the dictator can make corrections in the text as he sees it on the screen, and then, when the text is correct, can transfer it to a storage device, such as a magnetic disc storage unit, for later transmission to a relatively high-speed typewriter, printer, photocomposing or other recording device. Such a recording device, can for example, produce finished, addressed letters typed automatically without human intervention between the dictation and the typing process.

Thus there is provided an opportunity for very

substantial savings in labor costs in the production of typed and printed matter, and frees office personnel from typing duties and makes them available for more productive tasks.

5 Furthermore, the system and method enables the dictator to improve the quality of the written word he produces because he is able to make the corrections virtually immediately after dictation instead of later, when he has forgotten certain

10 matters which might require correction.

One of the potential problems with devices such as those described above is that many translators which are commercially available at the present time are relatively slow in operation.

15 Some of these machines cannot keep up with the dictator; that is, the translators cannot translate the words as fast as the dictator can dictate them. This is very unsatisfactory because it is inefficient, and may distract the dictator and cause him to

20 lose his train of thought.

Accordingly, one of the additional objects is to solve the foregoing problem and to provide a dictation system of the above described type in which the dictator need not wait for the translator to complete its translation before dictating the

25 next word.

This object is met, in accordance with one embodiment of the present invention, by the use of a audio recorder to record the dictation, and

30 control means for causing the recorded dictation be read out to the translator at a rate at which the translator can process it. Preferably the recorder also records tone signals which are used to control the translator and word processor. It is

35 preferred to use a recorder of the type in which dictation can be recorded at the same time that earlier dictation is being reproduced, and in which the dictation and reproduction can proceed simultaneously at two different rates.

40 Preferably, the reproducer is stopped when the pause between words is detected, and is re-started when the translator is ready.

This feature has an additional benefit in that the separation of the words from one another is

45 more distinct. This helps the translator to distinguish the words from one another and improves the accuracy of the translation. If the translator falls behind the dictator by any significant amount of time, the dictator can take

50 the opportunity to re-record portions of his dictation, if he wishes to change it.

A problem with all known dictating systems is that the preparation and handling of drafts is relatively inefficient. The time required in

55 delivering the draft to the dictator, hand-correcting the draft, and retrieving the draft from the dictator would be better spent on other tasks. Moreover, when the dictator is very remote from the transcribing station, it often takes a very long

60 time for the dictator to receive a draft of his dictation, correct it, and send it back for final typing. In fact, if the draft must be sent by mail, this can take a matter of days or weeks.

Accordingly, it is another object to provide a

65 dictation and transcribing system in which a draft

(or the final text) of the dictation is made available very quickly for the dictator to review, without the need for physical delivery of a copy of the text, and without regard to the actual distance of the dictator from the transcription station. It is a

70 further object to provide such a system and method in which corrections can be made very rapidly and easily, without writing them by hand. It is yet another object to provide such a system

75 and method in which the dictation and/or corrections are transcribed automatically by an automatic speech recognition device or "translator".

These objects are met by the provision of a

80 dictation system and method in which a dictation device and a visual display device are located at a dictation station. A transcriber and means for converting the transcribed words into coded electrical signals capable of being displayed on

85 the visual display device are provided at a transcribing station. Preferably, the devices at the dictation and transcription stations are linked for communication by means of either a directly-wired connection, or a telephone line, or a radio

90 line, or by other communication means.

The dictator dictates into the dictating device at the dictation station, and his dictation is transmitted to the transcription station where it is transcribed and converted into coded signals

95 which then are transmitted to the visual display device at the dictation station. The dictator then reviews the text on the visual display device, makes any necessary corrections, either by dictating them into his dictation device, or by making them electronically on his visual display

100 device, and then transmits the corrections to the transcription station. The corrections then are made at the transcription station and the text is typed in final form.

105 The dictator also can review the final text by means of the visual display device, and can make any further changes which may be necessary.

The transcription can be done either by an operator on a word-processing machine or

110 system, or by an automatic transcription device such as the device described above. In either case, the final text is printed automatically by the printer of the word processing system.

By means of the foregoing system and method,

115 the dictator can review a draft of his dictation much more quickly than if a typed or printed draft were prepared and carried into his office. In fact, he can review the text while it is still being transcribed. This facilitates more efficient

120 dictation, since the dictator has less chance to forget important corrections which are to be made. What is more, the dictator can make corrections either by dictation or electronically, instead of by hand. In most instances, either

125 dictation or electronic correction is faster than correction by hand. This invention is extremely advantageous when the dictator is remote from the transcription station, because it provides a means for the dictator to review the dictation

130 minutes, hours or even days earlier than if he had

to wait for a typed draft to be delivered to him.

In order that the invention may be more readily understood, various embodiments will now be described with reference to the accompanying drawings, in which:—

Figure 1 is a perspective view of a preferred dictation terminal as it sits on the desk of a dictator;

Figure 2 is a schematic circuit diagram of a system of which the dictation terminal of Figure 1 is a part;

Figure 3 is a schematic diagram of an alternative embodiment of the invention;

Figure 4 is an elevation view of a hand microphone and control unit of the dictation terminal shown in Figures 1 and 2.

Figure 5 is a side elevation view of the microphone shown in Figure 4;

Figure 6 is an elevation view of an alternative hand microphone of the type shown in Figures 4 and 5;

Figure 7 is a perspective schematic view of a complete automatic dictation typing system which might be used in the dictator's office;

Figure 8 is a schematic circuit diagram illustrating a system utilizing a portion of the apparatus shown in Figure 7 with several different dictation stations;

Figure 9 is a schematic circuit diagram showing the detailed interconnections of a portion of the system shown in Figure 2;

Figure 10 is a schematic circuit diagram of another system constructed in accordance with the present invention;

Figure 11 is a schematic circuit diagram of yet another embodiment of the system of the present invention;

Figure 12 is a partially schematic view of a visual display device utilized in the system of Figure 11; and

Figure 13 is a schematic representation of controls on the operation panel of the unit shown in Figure 12.

45 General Description

Figure 1 shows a dictation terminal 10 resting on the top of a desk 14. The dictation terminal 10 has a hand microphone 12, a video screen 16 for displaying written words which have been dictated, and a keyboard 18. Preferably, the unit 10 is of the type used in word processing systems, with the addition of the hand microphone 12, or a desk-top microphone and a foot pedal (not shown), if desired.

Figure 2 is a schematic diagram showing how the dictation terminal 10 is connected to the microphone 12 and other equipment in the dictation writing system. The unit 10 is one of four different remote units, each of which can be located in a different office or area, either in the same or another place of business.

Each of the units 10 is connected to one channel of a central translator unit 30 which translates the words dictated into the microphone

12 into binary digital data, and returns that data to the unit 10 which displays it on its video screen 16.

The video screen 16 preferably is capable of storing a full page of written text. The dictator can see each word as it is dictated and appears almost immediately on a screen 16. When the text on the screen is satisfactory, the dictator uses the keyboard to transfer the data for the page appearing on the screen 16 into a central disc file 32 which is of the type used with typical word processing systems. The unit 32 is a multi-disc magnetic disc storage unit.

The dictator then proceeds to dictate another screen full of written information, and transfers it to the disc file 32.

When the document being dictated is complete, the operator can operate the unit 10 to transmit the information stored in the disc file 32 to a printer 34, 36 or 38.

85 The Word Processor

Various word processing systems are suitable for use with the present invention. One such system is the "Dual Display" word processor sold by Dictaphone Corporation, Rye, New York.

Another is the "CPT 8000" word processor sold by CPT Corporation, Minneapolis, Minnesota. Such systems include video terminals with keyboards, disc files and printers, as it is well known. The above-identified word processors actually may have more features than would be necessary to make them suitable for use in this invention. Therefore, even simpler machines can be used, if desired.

The Printer

Although many different types of computer data printers can be used, for general office work, it is preferred that the printer be one like that which is used in most word processors, namely, a "daisy-wheel" printer or the equivalent, which produces typewritten matter on sheets of paper. Such sheets of paper can be letterhead paper or the like, so that the result of the printing operation is a typed letter, ready to be reviewed, signed and mailed. Additional printers 36 and 38 optionally can be provided with different types or sizes of paper so as to facilitate automatic typing on a variety of different media, merely by selecting the printer to be used. Alternatively, the additional printers can be used so as to enable two or more printing jobs to proceed simultaneously, thus increasing the production rate of the system.

The typewriter can be a standard typewriter or one of the proportional spacing type.

Alternatively, the unit to which the information is delivered can be a photocomposing machine which produces photographic film or paper upon which the written matter is recorded in order to be used in making printing plates.

If desired, the matter stored in each terminal 10 for printing can be delivered directly to the printer 34 instead of to the disc file 32.

The Translator

The translator 30 preferably is a commercially-available unit which is used for converting spoken words into coded electrical signals. Preferably, the device 30 should have a relatively large vocabulary. The device also should have the capability of handling inputs from several different sources simultaneously, if several dictators are to use the system. There are several devices which meet these requirements. For example, one such device is the Model VDES automatic speech recognition system sold by Interstate Electronics, Inc., Anaheim, California. This device has a vocabulary of up to 800 words, and can handle inputs from four different users simultaneously. It is believed that, in actual tests which have been performed, trained operators have achieved a recognition accuracy of over 99%.

Where only one dictator will use the system, and where a smaller vocabulary is acceptable, a lower cost translator can be used. For example, it may be possible to use the translator unit sold by Heuristics, Inc. of Sunnyvale, California. It is called the Model H2000 "Speechline" Automatic Speech Recognition system.

A number of other commercial systems are available which are believed to be satisfactory for use in the present invention. For their various capabilities and cost, see the article entitled "Words Into Action: I" by Gadi Kaplan, "IEEE Spectrum", June 1980, pages 22—26.

With most of the available translator devices, the speaker must pause briefly between successive words. This is because the machine is not capable of differentiating between words unless there is a certain minimum amount of time between them. However, some devices, such as the DP-100 device manufactured by Nippon Electric Co. Ltd., Tokyo, Japan, are capable of the limited recognition of "connected speech"; that is, words spoken without pauses between them, such as normally is done in ordinary speech. If the capability of recognizing connected speech is important, then such a machine should be selected. Another machine which is reportedly capable of detecting and recognizing connected speech is called the "Quiktalk" High-speed Speech Recognition System sold by Threshold Technology, Inc., Delran, New Jersey.

"Speaker independent" devices, that is, devices which recognize speech without programming for each individual user, are available. For example, such a device is sold by Dialog Systems, Inc. of Belmont, Massachusetts. The use of such devices will reduce programming time requirements, but may give reduced recognition accuracy, reduced vocabulary, and may be more costly than other systems.

The sales and repair literature, and the other published information concerning the above-described translators and other equipment discussed in the above-identified article by Gadi Kaplan, hereby is incorporated herein by reference.

Most translators must be programmed for

operation by a particular individual. The individual speaks each word of the machine's vocabulary a plurality of times. The machine then derives a pattern for the average of the signals received when the word is spoken, and stores this pattern in memory. Then, when the same person speaks that word during operation of the machine, the machine compares the incoming speech patterns with that stored and issues a code representative of the correct word.

Since the translator recognizes words strictly by their sounds, it cannot differentiate between homonyms. For example, if the dictator were to dictate the word "see", the translator could give the code representing "see", or the word "sea", or the letter "C". Which of these alternatives it would select depends upon how it is programmed. However, two of the three choices would be erroneous. A human operator usually can determine the proper spelling because of the meaning of the word as used in the sentence. Even then, it may be necessary for the dictator to give special instructions to avoid errors. It is believed that, at the present time, available translators are not capable of automatically differentiating between different spellings of the same sound.

Although a translator can be programmed to recognize proper names, names of cities, and towns and countries, etc., ordinarily it is impractical to program it to recognize more than a few frequently-used names because of the memory requirements and the programming time required. Although this may not provide an impediment for most of the current commercial uses of such translators, such as in quality control, etc., it creates a substantial impediment to the use of the translator in a dictation system.

A further problem of such translator devices is their relatively limited vocabularies. The largest vocabulary claimed for any of the commercial devices presently available is somewhat over 1,000 words. Were the dictator able to use only words in such a vocabulary, the machine probably would be of extremely limited usefulness and probably would be of little commercial interest as a dictating machine.

An additional problem is created when the vocabulary of the machine is made very large.

Since the translator usually must be programmed to the specific voice of a particular person, every person who uses the machine must repeat every word to be stored in the vocabulary during programming several times over. Therefore, the larger the vocabulary of the machine, the longer the operator must spend in initially programming the machine. Of course, the larger vocabulary makes the machine considerably more expensive, too.

125 The "Spell" Mode

In accordance with one aspect of the present invention, the foregoing problems are solved or alleviated by providing means whereby the dictator can switch the machine from its normal

mode into a "spell" mode in which each word can be spelled-out orally. The translator will recognize each character uttered by the dictator, and will assemble the characters together to form a word.

- 5 The machine also assembles that word together with other words previously or subsequently dictated, in order to form sentences. Further, if, for any reason, the use of the "spell" option is undesirable or unsatisfactory, the machine can be
10 operated in a "type" mode in which the words can be typed on the keyboard 18 of the unit 10, as in the normal operation of any word processor. However, it is believed that there will be little or
15 with the result that virtually all of the text is entered orally, rather than manually.

The Microphone

- Figure 4 is an elevation view of the hand microphone 12 used with the dictation unit 10 of
20 Figures 1 and 2. The microphone 12 includes a body or housing 42 having a grill 40 protecting a microphone inside the housing, and a cord 44 to transmit signals to and from the microphone.

- A plurality of function keys is located on the
25 microphone body. One such key is a "Dictate" key 48 which is pressed when the dictator wishes to have the word translated by the translator device.

- As it is shown in Figure 9, depression of the
30 "Dictate" key 48 sends a signal over a line 98 to the unit 10. The unit 10 is adapted so that when it receives a signal on line 98, it switches from operation with input from the keyboard 18, to operation with signals coming from the translator
35 processor is inhibited and it is adapted to receive and process codes from the translator.

- Also provided is a "Spell" button 50. On depression of the button 50, as it is shown in
40 Figure 9, a signal is sent over a line 100 to the unit 10. The signal also adapts the word processor to receive input signals from the translator, as in the "Dictate" mode. Also, the automatic word spacing provided in the "Dictate" mode is altered so that the characters are written
45 without spacing between them. Thus, the characters are assembled to form words.

- Preferably, the unit 10 is adapted to recognize the receipt on line 100 of a positive-going electrical pulse as an instruction to create an
50 interword space in the character train being recorded. If desired, a flip-flop circuit 102 can be connected in the manner shown so as to produce a positive-going pulse upon the release of the
55 "Spell" button 50 so as to create an inter-character space. Thus, the release of the "Spell" button at the end of each word which has been spelled will produce a space between that word and the next one.

- A third button on the microphone 12 shown in
60 Figure 4 is a "Type" button 52 which is depressed to change the system into the third mode of operation, namely, one in which input is from the keyboard of the word processor.

A number of other control buttons appear on

- 65 the handset 12. These include a "Backspace" button 54 to backspace by one character space; a "Word Back" button 56 to go back one word (that is, to the preceding interword space); and a button 58 to backspace by one entire line.

- 70 Also provided are buttons 60 to space in the forward direction by one character space; a button 62 to space one line forward, and a button 64 to space one word forward (that is, to the next interword space, when reviewing existing text).

- 75 Also provided are a button 66 to delete a character; a button 68 to delete an entire word, and button 70 to delete an entire line, all for the purposes of making corrections. These buttons actuate the correction mechanisms of the word
80 processor to make the corrections in a known manner.

- It can be seen from Figure 5 that the tops of the buttons are at different elevations from one another so as to make them easier to touch
85 without interference with adjacent buttons.

- A modified hand microphone unit 12 is shown in Figure 6. The unit shown in Figure 6 has a "Dictate" button 48 and a "Spell" button 50, as in Figure 4, and has several other buttons which
90 also appear in the device of Figure 4, and which are given corresponding reference numerals. The "Type" button 52 of Figure 4 has been omitted because it is not necessary. When the microphone of Figure 6 is used, the machine
95 automatically is in the "Type" mode at all times, unless one of the control buttons is pressed to change the mode of operation.

- Additional buttons which are provided in the device of Figure 6 include a "Paragraph" button
100 72, which causes the text to automatically shift to a new line and indent to start a new paragraph. Additionally, a button 74 is provided to capitalize words or characters being dictated.

- Also provided in the device of Figure 6 are a
105 "Store Word" button 76 and a "Recall" word 78. When a particular word, numeral or expression has been spelled out, it may be desirable to store the word in memory for a later recall so that the same word will not have to be spelled out several
110 different times while dictating a single text. The depression of button 78 causes the display of all the words stored in this manner on the screen of the unit 10, and allows selection of the desired word by operation of the keyboard 18. It is
115 desirable to locate as many of the function keys of the keyboard 18 on the microphone hand set as is practical.

Example

- As an example of the operation of the
120 foregoing system, the dictation steps which would be required for dictating the following letter will be explained. The letter is dated April 20, 1980, and is addressed to Mr. Joseph Jones, at Jones Men's Fashions, 4939 Hillside Avenue, Milwaukee, Wisconsin 53202.

Dear Mr. Jones:

We have received your letter of April 19th and your Purchase Order 4259 for three gross of

men's sheepskin caps at \$10.00 each for a total of \$4,320.00.

- 5 Please be advised that the price on this item now is \$11.50 each. If this price is acceptable to you, please sign a copy of this letter and return it to us to confirm your order at the new price.

Sincerely yours,

Stanley A. Penn
Sales Manager

- 10 The dictator starts by pressing the "Space Forward" button 60 (Figure 4) to properly locate the date. Then he presses the dictate button 48 and dictates the date in its entirety. The translator is programmed to correctly recognize the months of the year, the year 1980, and all numbers from 0 to at least 31, and it is programmed to cause a comma to be printed when the dictator dictates the word "comma".

- 20 Next, the dictator presses the "Line-Forward" button 62 to prepare for the dictation of the address of the letter. Since the name and address of the addressee are not easily programmable in the automatic dictation mode of the machine, the operator now depresses the "Spell" button 50 to enable him to spell the name and address of the addressee. He then orally spells the addressee's name. Capital letters preferably are formed by pressing the "CAP" button 74 (Figure 6). However, they also can be formed by programming the machine to capitalize the next character when a code word is spoken. The dictator then presses the "Line-Forward" button 62 and orally spells the address of the addressee. He then presses the "Line-Forward" button again, releases the "Spell" button, depresses the "Dictate" button to dictate the whole words: "Dear Mr.". (The machine is programmed to spell "Mr." when it hears "mister".) Then the dictator reverts to the "Spell" mode to spell "Jones", and dictates the word "colon" to produce the colon. The dictator then proceeds to dictate the text of the letter.

- 45 The machine has sufficient vocabulary to translate the first sentence of the text up to the number "4259". Since the numbers 4259 normally would be printed with spaces between them, the operator depresses the "Spell" button 50 and dictates the numbers "4259". The machine is not able to correctly translate "for", since it has homonyms. For example, the same sound could mean the numeral "4" or "four" as well as the word "for". Therefore, the dictator shifts into the "Spell" mode and spells out the word "for". He then shifts back into the dictate mode by depressing the button 48 until he reaches the expression "men's sheepskin caps". Since these words are not found in the vocabulary of the machine, the dictator switches to the "Spell" mode and spells these words orally, releasing the "Spell" button 50 at the end of each word in order to space the words from one another. Similarly, when the word "for" and the

sum "\$4,320.00" are reached, these words and numbers also are spelled out.

- 65 The operator then depresses the paragraph key 72 (Figure 6), if such a key is provided, which automatically spaces forward one line and indents to the start of a line. Alternatively, he can say "paragraph", or give another oral command, and the machine, when specially programmed to do so, will space and indent automatically. The remainder of the letter, except for the \$11.50 price, can be translated by the translator unit 30, so that the "Spell" mode need be used only once more during the dictation of the letter. Although the name of the writer, Stanley Penn, is a proper name, it is stored in memory so that the translator can recognize it, because it will be used repeatedly by Mr. Penn and this makes it worthwhile to store the name.

- 80 The depression of the "Capital" button, either on the keyboard 18 or on the hand microphone 12, automatically capitalizes only the first letter of the word being capitalized. If the capitalization of every letter of the word is desired, this can be accomplished by holding the "Capital" button down while continuing dictation.

An Individual System

- Figure 7 shows a complete individual dictation system 80 which might be used in a one-man office, or by a person in a larger office desiring to have all the equipment nearby. The system 80 includes a separate translator unit 82 with a hand microphone 12. The translator 82 is connected to the video keyboard unit 10, which is connected to the disc file 32 and the printer 34. Paper can be fed into the printer 34, and the printed or typed text can be taken out of the printer by the dictator himself.

- 100 Figure 8 shows a multiple-terminal dictation system using a plurality of devices 10 and individual translators 82 at different work stations 86, 88, 90 and 92, but using a single disc-file 32 and printer 34.

- 105 At each station, the output of the microphone 12 is delivered to the translator 82, which delivers its output to a switching device 94 which switches the unit 10 between the typing mode, the "Spell" mode, and the "Dictate" mode, and operation is substantially as described above. The output of each station is transmitted by means of a multiplexer circuit 96 to the disc file unit 32. The multiplexer unit 96 may contain a buffer storage device to store data received from one of the dictators until later when it can be recorded in the disc file. This enables simultaneous operation of the various work stations without interference between them.

Paper Handling

- 120 It is possible that the system of either Figure 2 or Figure 8 can be used with a single printer which is tended by an operator who puts in the desired sizes of paper, takes out the typed products, puts them in envelopes, and mails them. However, if the system of Figure 7 is used,

it is desirable that the paper be fed from a roll so that the repeated insertion of sheets by the dictator is not required. If letterhead paper is desired, the letterhead can be printed at regular intervals along the sheet. These headings are printed at distances from one another which are excessive for normal use, and the excess is trimmed off by a knife or cutter (not shown). If a particular page is to be unadorned or blank, the heading simply can be cut off of the top of the next sheet, etc.

If desired, an automatic sheet feeding device can be provided as an input to the printer 34 in order to input 8-1/2" by 11" letterhead, plain bond, or larger sizes of paper, as desired. Alternatively, each of the separate printers 34, 36 and 38 can be set up to print on a specific type of roll fed paper, with the paper being cut exactly to the desired length by a knife in the machine. Selection of the printer determines the type of paper used. This can be done electrically by the dictator.

If desired, the operation of the machine can be set up in different formats by the use of conventional computerized format control. Thus, by simple selection of the desired format by operation of the keyboard 18, the machine can be adapted automatically to set up for the preparation of letters, or reports, etc. This can be done in accordance with techniques well known in the art.

Remote Communications Embodiment

Figure 3 shows an alternative form of the invention in which the input to the device is by means of a telephone handset 106. The handset 106 is connected through the telephone lines 107 remotely to a receiver 109 at the location of the dictation writing device. The translator 82 is pre-programmed to recognize only a specific caller's voice. When it does so recognize his voice, it converts the words used by the speaker into digital form, and sends them to a speech reproduction device 106 which transmits audible reproductions of the words over the telephone lines back to the dictator in order to allow him to check the correctness of the translation. After initial identification, the dictator can dictate remotely and have his words stored in the disc-file, and then he can cause the transmission from the disc-file of the dictation to a printer 34. Thus, a remote dictation feature has been provided for the invention; one which requires no separate hand-held code sending device for remote actuation.

If preferred, the telephone handset and lines can be replaced by a radio transceiver, or by other types of remote communication devices.

All function instructions such as "spell" and "cap", etc. can be spoken and need not be input by means of pushbuttons, in this embodiment of the invention. This is accomplished by programming.

Similarly, the translator 82 can be used for remote identification of a subscriber or owner of a

telephone answering machine 108. When the caller has been properly identified, the telephone answering machine 108 will automatically read the messages stored in the machine out to the caller and allow him to give the machine new instructions. Thus, the invention provides for the remote retrieval of information from an automatic telephone answering machine, without the usual hand-held coding device.

One use envisioned for the present invention is in ordinary offices. Another is for use as an input device to type composing machines, particularly phototypesetting machines. Such a system can be used, for example, in composing a daily newspaper. Each reporter can dictate his story at a terminal of the type shown in Figure 1, and he can review and edit the column before it actually is composed by the photocomposing machine.

One of the advantages of the invention is that the dictator sees the product of his dictation on the video screen 16 virtually immediately after he has dictated it. This gives him an opportunity to edit or correct the text while the subject matter of the dictation is fresh in his memory. Furthermore, since the letter or other document is typed virtually immediately after dictation is complete, the dictator can see a copy of the typed or printed text shortly after it has been dictated, thus avoiding the often substantial delay in transcription of dictation by usual means.

It is believed that the foregoing factors may be enough to improve the overall dictation efficiency of the dictator, compared with his efficiency when using other dictation equipment. When this is coupled with the ability to mail letters more promptly and to otherwise complete tasks at an earlier date, the overall improvement should substantially outweigh any increase in the time of dictation required to spell selected portions of the dictation. In this regard, it should be noted that in normal dictation it often is required that the dictator spell certain unusual words, names, addresses, towns, etc. Therefore, the increased amount of time required to spell additional words not capable of being translated correctly is not as great as it otherwise might be.

As translator devices improve with further development, it is probable that the size of the available vocabulary will increase without a corresponding increase in cost, and the programming time will decrease, so that progressively fewer words and terms must be spelled.

Another time-saving feature of the invention is provided by the fact that the dictator need not operate a rewind mechanism and hunt for previously dictated material because this material normally will be in full view. Thus, if he forgets what he said previously, he merely need to refer to the screen quickly, without operating any buttons, to regain his train of thought. If the material he is looking for is on a previous page, most word processing machines have the capability of recalling the previous page or an earlier page of text rapidly and easily.

In a preferred embodiment of the invention, it is preferred that all characters, punctuation marks and numbers be dictated only in the "Spell" mode, and that certain homonyms of those items can be programmed to be retrieved during the "Dictate" mode. For example, during operation in the "Dictate" mode, dictation of the sound for the letter "r" would be translated as "are". However, during the "Spell" mode, the same sound will be translated as the letter "r" (or "R"). Similarly, during the "Dictate" mode, the sound for the numeral "2" would be translated as "too" (or "two", if preferred). However, during the "Spell" mode the same sound would be translated as the numeral "2".

Of course, all punctuation marks either must be dictated or inserted by means of keys on the microphone 12 or the keyboard 18. During the "Dictate" mode, a "." punctuation mark would be translated as the word "period". However, during the "Spell" mode, the same sound is translated as ".".

Capital letters also can be handled either by the use of a key on the microphone 12, or on the keyboard 18, or by dictation of the word "Capital" preceding each letter to be capitalized. The machine can be programmed to automatically capitalize the first letter of the first word of each new sentence. Similarly, it should be programmed to automatically provide for two inter-word spaces following each period punctuation mark.

By means of the foregoing separation of letters and numbers into the "Spell" mode only, a form of automatic treatment of certain homonyms has been accomplished.

Accelerated Dictation System and Method

Figure 10 shows a dictation and transcription system 110 which is substantially the same as the system of Figure 1, except that the system 110 includes a recorder/reproducer unit 16 which provides an improvement in the operation of the system.

The system of Figure 1 comprises a plurality of dictation devices, each of which includes a visual display device 120 and a microphone 112. The dictation spoken into the microphone 112 is delivered to a central translator device 118 which converts the spoken words into electrical signals representing the corresponding written words, and those signals are delivered to the visual display device 120 where the words are displayed on the screen for the dictator to see.

The dictator can correct the text which he sees, and then deliver it to a disc file 122 or other storage device for storage, and thence to a printer or a photocomposer 124 or 126 to prepare a printed text. There are two dictation stations shown at Figure 10. Each dictation station is in a different office or location in one or more buildings, and the translator unit 118, disc file 122 and the printer 124 or 126 can be located at a central location within the same building or a different building. Although only two dictation stations are shown in Figure 10, it should be

understood that this has been done solely to simplify the drawings, and that more dictation stations can be used, if desired.

As it is explained in greater detail above, special control means are provided to enable the dictator to selectively spell words which the translator is incapable of correctly translating, so as to avoid the need for a very large vocabulary for the translator, and to overcome other shortcomings of that device. Pushbuttons 114 on the microphone 114 are used to create signals indicating the selection of the spelling mode, as well as to delete, backspace, etc., in order to make corrections in the text appearing on the screen of the CRT 120.

In accordance with the present invention, a sound recorder/reproducer device 116 is interposed between the microphone 112 and the remainder of the system. Preferably, the recorder/reproducer 116 is of the type in which recording and reproduction can take place simultaneously and at different rates. One device which has the capability, for example, is an endless-tape random-storage recorder/reproducer sold under the trademark "Thought Tank" by Dictaphone Corporation, Rye, New York. That device has a storage housing 128. In the housing is an endless magnetic tape 136 which is "jumble-stored" (allowed to pile up randomly in the housing), a recording unit 130 and a reproducing unit 134. A separate capstan 132 driven by its own motor moves the tape 136 past the recording head of the unit 130, and another capstan and motor moves the tape past the reproducing head of the unit 134. With such a device, dictation can be reproduced within a few seconds after it has been dictated, and reproduction can take place simultaneously with and independently from dictation. Moreover, the reproduction and dictation rates can be quite different. If the transcription lags behind the dictation, the tape bearing the dictation will accumulate in the housing 128 until it can be transcribed.

Voice signals are delivered over a line 113 to the recording unit 130 which records them on the tape 136. Signals indicating the selection of the "spell" mode of operation of the translator also are delivered over the line 113 and recorded on the tape. These signals preferably are audio-frequency tone-coded signals developed by tone generators operated by the push-buttons 114, in the nature of "Touch-Tone" telephone pushbuttons. Other signals to be used in the operation of the translator similarly are recorded on the tape. Decoding circuitry is provided in the translator to decode the tone coded signals and instruct the translator in its operations when the tones are reproduced by the reproducing unit 134.

Also transmitted over the line 113 are signals which are used to stop, start and control the recording unit 130 as in the normal operation of the recorder/reproducer 116.

Other signals developed by operation of the

pushbuttons 114 in order to control the operation of the visual display unit 120 for corrections, etc., are delivered directly to the unit 120 over a line 115.

5 At the start of operations, a START signal is delivered through an OR gate 141 to the reproducing unit 134 to start it. Preferably, the START signal is developed by the operation of a key on the keyboard of the display unit 120.

10 Connected to the reproducing unit 134 is a detector circuit 138 which detects the gaps between words or function signals (pauses of at least 0.1 second duration which are required between successive words in the dictation for proper operation of the translator) and disables the reproduction device 134 and stops movement of the tape past the recording head. This condition persists until another detector circuit 140 detects a signal sent over a line 117 upon the delivery to the visual display unit 120 of the coded signals representing a word which has been translated by the translator 118 and sends a signal through the OR gate 141 which starts the reproducing unit 134 again. The circuit 140 also sends a reset signal to the detector 138 to reset it and ready it for the detection of the next inter-word gap. The detector 138 also sends a signal over a line 119 when it detects a gap so as to indicate the end of each word more positively than if the gap were detected solely by the translator.

Thus, by means of the foregoing construction, the recorded words are reproduced one-at-a-time, at a rate at which the translator is capable of translating them. If the translator is finished with the translation of a word before the next inter-word gap is detected, the reset signal from the detector 140 will prevent the circuit 138 from stopping the tape, and the next word will be reproduced without stopping the reproducing unit.

40 In addition to, or instead of, the recorder/reproducer 116, a digital buffer storage unit can be used to store voice and translator function signals.

45 Circuits for detecting a pause in voice signals and turning a device on or off in response to such a detection are known and used, for example, in automatic telephone answering machines, and they will not be described in detail herein.

50 The above-described system and method allow the dictator to dictate at his own pace, without regard to whether the translator can keep up with him. The translator proceeds at its own pace, utilizing the recorded dictation as fast as it can. It is believed that this system and method take advantage of the pauses which a dictator normally has in his dictation. That is, if the translator lags behind, during the pauses which normally occur in the dictation, the translator continues to translate stored dictation, thus enabling it to utilize the time which otherwise would be wasted to help to match the speed of the translator to that of the dictator.

60 It also is believed that the stop-start operation of the reproducer helps to enhance the correct

detection of the gaps between words. This is because the reproducer actually stops between words. This is believed to improve the correct translation of the words being dictated.

70 Remote Dictation System and Method

Figure 11 describes a remote dictation system including two dictation stations 142 and 144, and a transcription station 148. The dictation stations 142 and 144 are in two different offices in one building 146, while the transcription station 148 is in another building. Of course, the areas 142, 144 and 148 also can represent different areas within the offices of a single business establishment. Moreover, although only two dictation stations 142 and 144 are disclosed, it should be understood that the system can include more dictation stations, if desired.

At each dictation station 142 or 144 there is a dictation device 178 or 180 and a visual display device 170 or 172. The dictation device 178 is shown as a telephone hand-set type of input device for a "Thought Tank" remote dictation system such as the one described above.

The "Thought Tank" dictation system includes a recorder/reproducer device 164 at the transcription station, as well as a head-set 168 for the operator to use in listening to the dictation. Dictation is transmitted over a line 162 to the remote recorder/reproducer 164.

95 The line 162 is shown schematically. It can represent either a wire extending between the offices of the dictators and the transcription station, or it can be a telephone line, or it can represent a radio, video, or other communication link suitable for transmitting dictation to the transcription station.

The visual display device 170 or 172 preferably is a CRT device of the type used in word processing systems. The visual display devices 170 and 172 are connected to the transcription station 148 by means of the same line 162.

Also located at the transcription station 148 is a word processing system generally indicated at 150. The word processing device 150 includes a CRT display unit 152 with a keyboard 158, a disc file 154, and a relatively high-speed printer or composer 156. The coded signals which form the output of the word processor are delivered over a line "A" or "B", etc., through the line 162 to the visual display device 170 or 172 which has the corresponding letter next to it.

In the case in which the line 162 is a telephone line, the word processor output is coupled to the line 162 by means of a modem 160, which converts the digital signals from the word processor into audio signals suitable for sending over telephone lines. Similarly, each of the visual display devices 170 and 172 is connected to the line 162 through its own modem 176 or 174, respectively. In this case, the input units 178 and 180 are telephone hand sets which are coupled to the telephone line 162 by means of normal telephone coupling devices 182. At the

transcription stations 148, a unit 166 is used to couple the recorder/reproducer 164 to the telephone line 162. The unit 166 is a standard unit which is used to couple the recorder/reproducer 164 for remote dictation, and will not be described in detail herein.

An automatic word recognition device or translator 169 is shown in dashed outline at the transcription station in Figure 11. It is shown in dashed outline to indicate that it can be used instead of an operator to automatically translate and thus transcribe dictation, in the manner described above.

The system shown in Figure 11 operates as follows.

The dictator dictates into the device 178 or 180, and this dictation is transmitted to the recorder/reproducer 164 where it is recorded. Shortly thereafter, or at a later time which is convenient, the operator reproduces the dictation and listens to it by means of the headset 168 (or the translator 169 translates the dictation).

In transcribing the dictation, the operator uses the keyboard 158 to produce the dictation and display it on the screen of the CRT display device 152, or the translator 169 performs the same function. Simultaneously, signals are transmitted over the line 162 to the visual display device 170. The words there are formed on the screen of the visual display device so that the dictator can review the dictation for making corrections or other purposes.

When making corrections, the dictator can dictate them into the input device 178 so that they are recorded in the recorder/reproducer. Alternatively, corrections can be made electronically, in the same manner in which they are made in word processor systems.

If the corrections are not made electronically, the operator listens to the corrections and makes them. In either case, after corrections have been made, the text is sent to the disc file and/or the printer 156 to produce the final printer copy. If desired, the dictator can again review the printed text after the corrections have been made and before the printed copy has been prepared.

Figure 12 is a schematic elevation view of one of the identical visual display devices 170 and 172. The unit 170 includes a CRT screen 184, and a control panel 186. The CRT screen 184 has a vertical array 188 of line markings in the left hand margin to facilitate oral reference to specific lines of the text by the dictator when dictating corrections.

The most basic components of the control panel are enclosed in Figure 13 in dashed outline 190. These controls include an on/off switch 192, a "page forward" switch 194 to change the display on the screen to the next page of text material, and a "page back" switch 196, to change the page of material back one page.

Also included in the controls 190 is an indicator light 198 indicating that transcription is in progress, and a second indicator light 200

for review. The indicator 198 is lighted whenever transcription is being taken from the recorder/reproducer 165. The indicator 200 is lighted by the operator or the translator device when a transcription job is complete, and can be extinguished by the dictator. Thus, the dictator can determine when transcription is in progress and when it is complete so that he can decide when to turn on the unit 170 to review the text.

Also shown in Figure 13 are optional controls. These include a keypad 202 including keys 204 for entering each of the numerals from 0 to 9, as well as keys 206 for moving a cursor up, down, to the left and to the right on the CRT screen 184.

This cursor is for the purpose of indicating the location at which a particular correction is to be made. Also included is an "Enter" switch 108 which enters the number selected on the keypad 202. Means are provided for transmitting the location of the cursor to the transcription station in response to operation of the Enter switch.

Additional keys include a "Display Document" key 210 which is used to display a particular document which has been stored in the word processor system, regardless of the order in which the documents there are stored. This key is used in conjunction with the keypad 202, which is used to identify the document desired by its identification number. The button 210 is pressed to cause the selected document to be displayed.

A "Document Forward" switch 12 and a "Document Back" switch 214 also are provided. These switches will allow the dictator to review one document stored immediately preceding the document being displayed, or a document following the document being displayed. The circuitry and programming necessary to perform the functions and operations controlled by these switches are well known in the art and will not be described in detail herein.

As an alternative to oral identification and cursor location of corrections, a "light pen 210 and control circuit 213 can be used to quickly identify the location of the correction and transmit that location to the transcribing station.

The system shown in Figures 11 through 13 has the advantages described above. It is possible for the dictator to see a draft of his dictation (or the final copy, if desired) on the visual display device very shortly after he has dictated the dictation. This enables him to review and correct it before he has forgotten what it is all about. This tends to make for greater dictation efficiency.

The draft is in front of the dictator more rapidly than it would be if it were a typed or printed draft. This is because the draft is delivered by electrical means rather than manually.

Another advantage is that paper for the draft is not wasted. The only paper which is used is that for the final copy. Another advantage is that corrections can be dictated, if desired, thus making the correction process potentially faster than if corrections were made by hand.

The system of Figure 11 makes it possible to provide a centralized transcription service for a

relatively large number of dictation stations which can be located on the same floor in the same building, or in separate buildings within an industrial complex, or in separate buildings within a city or municipal area, or even in different cities. This permits the provision of transcription services for even the most remote outposts, where such services normally would be totally impractical to provide. Moreover, it facilitates a transcription service operation in which dictation is transcribed by the independent agency operating at the transcription station, and subscribers are connected by telephone or other communication links to the transcription service so that fast transcription can be provided without the physical delivery of typed copies to and from the transcription service offices.

The above description of the invention is intended to be illustrative and not limiting.

Various changes or modifications in the embodiments described may occur to those skilled in the art and these can be made without departing from the scope of the invention.

Claims

1. A device for converting speech into corresponding written words, said device comprising, in combination, transducer means for converting speech into electrical signals, translator means for converting signals from said transducer into coded signals representing words and characters of the alphabet, visual display means adjacent said transducer means for receiving said coded signals and displaying the corresponding words and characters, for selectively forming said characters into other words, and for assembling said other words with the first-named words to form written sentences.

2. A device as claimed in claim 1, including storage means for storing said electrical speech signals prior to being delivered to said translator means, and means for delivering said speech signals from said storage means to said translator means at a rate at which said translator means is capable of translating them.

3. A device as claimed in claim 1, or 2, including correction means for deleting and correcting words and characters displayed by said visual display means.

4. A device as claimed in claim 3, in which said correction means includes said translator for developing replacement words and characters.

5. A device as claimed in any preceding claim, in which said visual display includes a cathode ray tube screen for displaying a full page of said words.

6. A device as claimed in any preceding claim, including graphic means for recording said words and sentences on sheet material.

7. A device as claimed in claim 6, in which said graphic means is a printer.

8. A device as claimed in claim 6, in which said graphic means is a type composing machine.

9. A device as claimed in any preceding claim,

including storage means for storing said coded signals for later recording on sheet material.

10. A device as claimed in any preceding claim, including manual control means for switching said visual display means between a word-receiving mode and a character-receiving mode, said display means being adapted, in the last-named mode, to form said character signals into said other words by selectively locating the characters adjacent one another until a spacing command is received.

11. A device as claimed in claim 10, in which said transducer means is a microphone, said manual control means being mounted integrally with said microphone.

12. A device as claimed in claim 10 or 11, including spacing means for automatically spacing each of said other words from any preceding words upon the operation of said manual control means, and providing proper spacing of each of said other words from the next following word.

13. A device as claimed in any preceding claim, including a keyboard and means for permitting character codes to be input to said visual display means from either said translator or said keyboard.

14. A device as claimed in claim 2, or claim 2 in combination with any of claims 3 to 13, in which said electrical speech signal-storage means comprises a voice signal recorder device capable of recording speech and simultaneously reproducing and delivering electrical signals representative of previously stored speech, the rate of recording and reproducing being independent of one another.

15. A device as claimed in claim 14, in which said recorder device is a random-storage endless magnetic tape recorder-reproducer.

16. A device as claimed in claim 2, or claim 2 in combination with any of claims 3 to 15, including sound reproducing means for reproducing dictation, said delivering means comprising means for stopping and starting said reproducing means as necessary.

17. A device as claimed in claim 16, including means for detecting pauses between words and causing said reproducing means to stop in response to the detection of said pauses.

18. A dictation device comprising: a dictation terminal consisting of a video display device for displaying dictated words, and a microphone for receiving dictation; translator means for converting signals from said microphone into coded electrical signals representing words or characters; means for delivering said signals to said video display device; means for enabling the dictator to audibly spell words not capable of correct translation by said translator means, and for assembling the latter words with others on said video display device; and printer means for printing the text so prepared.

19. A method of converting speech into written form utilizing a speech recognition device which translates spoken word and character sounds into

corresponding coded electrical signals, said method comprising the steps of; speaking into said device those words which are translatable by said device, speaking into said device characters which spell words which are not translatable by said device, selectively assembling the resulting character signals to form word signals, and visibly displaying words corresponding to said word signals.

20. A method as claimed in claim 19, in which said displaying step comprises displaying said words on a video screen, and selectively correcting said words electronically.

21. A method as claimed in claim 19 or 20, including recording said words on sheet material, said recording step being selected from the group comprising printing and photocomposing.

22. A method as claimed in claim 21, including storing said word signals in memory prior to recording them.

23. A method as claimed in any of claims 19 to 22, including the step of assigning one spelling of homonyms to one mode of operation of said device, and another spelling to the "spelling" mode of operation, whereby a predetermined differentiation between said spellings is made.

24. A method as claimed in claim 19, which includes storing speech signals prior to delivering them to said speech recognition device, and reading said speech signals out of storage at a rate corresponding to the rate at which the signals can be translated by said speech recognition device.

25. A method as claimed in claim 24, in which said speech signal-storing step is performed by recording said signals in a recorder/reproducer device, and stopping and starting the reproduction function of said recorder/reproducer device when necessary to allow said speech recognition device to complete the translation of words in process.

26. A method as claimed in claim 25, in which said stopping step comprises detecting pauses between words and stopping the reproduction function in response to such detection.

27. A remote dictation system comprising a receiver for receiving remote voice signals, a translator unit programmed to receive and translate the speech of a specific individual, and printer means for printing words corresponding to said speech, said receiver delivering remote voice signals to said translator.

28. A system as claimed in claim 27, including means for transmitting back to the sender an audible reproduction of the output of said translator device.

29. A remote retrieval device for messages in an automatic telephone answering device, said device comprising a translator unit programmed to receive and translate the speech of a specific individual, and means for enabling the transmission of the message stored in the telephone answering device to the remote caller whose voice matches the stored data for the individual caller.

30. A dictation system comprising, in combination, at least one dictation device and a visual display device located at a dictation station, transcription equipment including, means operable for converting the dictated words into electrical signals capable of being transmitted to said display device and displayed as visible words on said display device, and graphic representation means for graphically converting said signals into written form.

31. A system as claimed in claim 30, in which said transcription equipment is located at a transcription station which is remote from said dictation station.

32. A system as claimed in claim 31, including telephone means for transmitting dictation from said dictation device to said reproducer means at said transcription station.

33. A system as claimed in claim 32, in which the means for transmitting dictation comprises radio transmission means, telephone transmission means, and/or direct wire transmission means.

34. A system as claimed in any of claims 30 to 33, in which said visual display device is a video monitor, and said graphic representation means is a relatively high-speed printer.

35. A system as claimed in any of claims 30 to 34, in which said visual display device at said dictation station includes means for indicating that transcription of dictation is in progress.

36. A system as claimed in any of claims 30 to 35, in which said visual display means includes means for indicating that transcription is complete.

37. A system as claimed in any of claims 30 to 36, in which said visual display means includes means for changing the displayed matter by a page forward or a page back.

38. A system as claimed in any of claims 30 to 37, in which said visual display device is a video monitor, and including means for moving a cursor on the video screen of said video monitor and transmitting the location of the cursor to said transcription station.

39. A system as claimed in any of claims 30 to 38, including means for selecting and identifying a particular document, displaying that document, and changing the document displayed from one to another selected document.

40. A system as claimed in any of claims 30 to 39, in which said transcription equipment includes sound reproducer means for reproducing dictation from said dictation device, and in which said dictation reproducer is a recorder/reproducer of the random-stored endless magnetic tape variety.

41. A system as claimed in any of claims 30 to 41, including a plurality of said dictation devices at a plurality of dictation stations.

42. A system as claimed in any of claims 30 to 42, in which said transcription equipment includes automatic speech recognition means for translating speech into electrically coded signals representative of words and characters and

delivering said coded signals to said visual display device.

43. A system as claimed in any of claims 30 to 42, including a light pen for indicating the locations of corrections, and means for transmitting said locations to said transcription equipment.

44. A method of dictation and transcription, comprising the step of dictating dictation at a first station, transcribing said dictation at a second station, converting said dictation into transmittable coded signals and transmitting said coded signals to a visual display device at said first station, making corrections in the text displayed at said first station, transmitting said corrections back to said second station, making said corrections at said second station, and preparing a written text of said dictation at said second station.

45. A method as claimed in claim 44, in which said first station is remote from said second station.

46. A method as claimed in claim 45, in which said dictation and corrections are transmitted electrically by a method from the group consisting of telephone, wireless, and wired transmission.

47. A method as claimed in claim 44, 45 or 46, in which said transcribing step is performed by an

automatic speech recognition system.

48. A device as claimed in any of claims 44 to 47, in which said dictation step is accomplished by the use of a telephone hand set delivering signals through telephone lines to a recorder/reproducer at said second station.

49. Devices for converting speech into corresponding written words, substantially as hereinbefore described with reference to the accompanying drawings.

50. Dictation devices or systems embodying devices as claimed in claim 49.

51. Dictation devices or systems, substantially as hereinbefore described with reference to the accompanying drawings.

52. Remote retrieval devices for messages in automatic telephone answering devices, substantially as hereinbefore described with reference to the accompanying drawings.

53. The methods of converting speech into written form, substantially as hereinbefore described with reference to the accompanying drawings.

54. Dictation and transcription methods embodying the methods as claimed in claim 53.

55. Dictation and transcription methods substantially as hereinbefore described with reference to the accompanying drawings.